
Forme et fonction de la périphérie gauche dans un corpus oral multigenres annoté

Laurence J. Martin, Liesbeth Degand et Anne-Catherine Simon



Édition électronique

URL : <http://journals.openedition.org/corpus/2509>

DOI : 10.4000/corpus.2509

ISSN : 1765-3126

Éditeur

Bases ; corpus et langage - UMR 6039

Édition imprimée

Date de publication : 1 décembre 2014

Pagination : 243-265

ISSN : 1638-9808

Référence électronique

Laurence J. Martin, Liesbeth Degand et Anne-Catherine Simon, « Forme et fonction de la périphérie gauche dans un corpus oral multigenres annoté », *Corpus* [En ligne], 13 | 2014, mis en ligne le 01 mai 2015, consulté le 08 septembre 2020. URL : <http://journals.openedition.org/corpus/2509> ; DOI : <https://doi.org/10.4000/corpus.2509>

Forme et fonction de la périphérie gauche dans un corpus oral multigenres annoté

Laurence J. MARTIN, Liesbeth DEGAND,
Anne-Catherine SIMON¹
Centre de recherche VALIBEL – Discours et Variation
Institut Langage et Communication
Université Catholique de Louvain,
Louvain-la-Neuve, Belgique

Introduction

Toute unité de discours, quels que soient le modèle discursif de référence et la méthode de segmentation utilisée, est délimitée par des frontières gauche et droite. L'hypothèse que nous poursuivons dans notre recherche est que nous pouvons décrire les périphéries gauches des unités de discours en termes de paradigme, celles-ci devant partager des propriétés fonctionnelles communes. Dans ce but, il convient, d'une part, de choisir une unité de segmentation du discours pour l'analyse et, d'autre part, de formellement identifier ces unités périphériques pour ensuite les analyser formellement et fonctionnellement.

Cette étude s'inscrit dans le cadre des recherches centrées sur la question de la corrélation entre les structures prosodique et syntaxique (voir notamment Lacheret-Dujour & Victorri, 2002 pour des premiers travaux sur le sujet). Les unités de découpage servant de point de départ à notre étude sont les Unités de Base du Discours (désormais BDU pour *Basic Discourse Unit*), définies par Degand & Simon (2008, 2009a, 2009b, 2011) comme « the segments that speakers use to build a representation (interpretation) of the discourse, i.e. a kind of 'minimal discourse interpretation segments' » (Degand & Simon, 2009a : 4).

¹ laurence.j.martin@uclouvain.be, liesbeth.degand@uclouvain.be,
anne-catherine.simon@uclouvain.be

Partant du postulat selon lequel ni la complétude syntaxique (Roulet, 1991) ni la complétude prosodique (Chafe, 1994) ne sont une condition suffisante pour la délimitation d'unités dans le flux de parole (Halford, 1996 et Selting, 2000), nous identifions les BDU selon une double analyse (largement détaillée dans Degand & Simon, 2009a), où unités syntaxiques et unités prosodiques sont extraites de manière totalement indépendante :

Les unités syntaxiques, d'une part, sont identifiées via une analyse basée sur la grammaire de dépendance (en particulier, Blanche-Benveniste *et al.*, 1990), qui résulte en une segmentation du flux de parole en unités de rection, celles-ci étant ensuite segmentées en séquences fonctionnelles (selon le découpage décrit par Bilger & Campione, 2002) ;

Les unités prosodiques, d'autre part, sont identifiées sur base des frontières prosodiques majeures, qui sont les seules retenues pour l'analyse. Celle-ci est réalisée de manière semi-automatique et se base sur la détection de syllabes proéminentes en fonction de leur hauteur mélodique, leur durée relative et la durée de la pause subséquente (Mertens & Simon, 2013).

Cette méthode d'analyse tient compte à la fois de la représentation statique de la structure du discours (unités syntaxiques et représentation sémantique) et de son déroulement dynamique (unités prosodiques). Les structures prosodique et syntaxique offrent chacune des indices linguistiques de surface pour la délimitation d'unités dans le flux de parole, et la combinaison de ces deux types d'unités permet de définir les unités à un niveau supérieur, le niveau discursif.

La mise en correspondance de ces deux segmentations est réalisée de la manière suivante : chaque fois qu'une frontière d'unité de rection coïncide avec une frontière prosodique majeure, une BDU est identifiée. Cela donne lieu à quatre types de BDU distinctes, représentés en figure 1 : (a) les BDU congruentes, où une unité de rection correspond à une unité prosodique majeure ; (b) les BDU groupées par l'intonation, où une unité prosodique regroupe plusieurs unités de rection ; (c) les BDU groupées par la syntaxe, où une unité de rection est segmentée en plusieurs unités prosodiques majeures ; et (d) les BDU ré-

Forme et fonction de la périphérie gauche

gulatives, constituées d'un élément syntaxiquement non régi (connecteur, adverbe de phrase ou marqueur discursif) et isolé par une frontière prosodique. Enfin, une dernière catégorie, étiquetée « BDU-x », regroupe les restes de l'analyse, c'est-à-dire les BDU regroupant plusieurs unités syntaxiques et plusieurs unités prosodiques majeures.

Syntaxe	[]	[]	[]	[]	<	>
Prosodie	[]	[]	[]	[]	[]
BDU	[BDU-C]		[BDU-I]		[BDU-S]		[BDU-R]	
	Congruente		groupée par		groupée par		régulative	
			les unités		les unités			
			prosodiques		syntaxiques			

Figure 1. Quatre types de BDU

En étudiant la distribution des BDU dans quatre genres de discours, Degand & Simon (2009b) ont constaté qu'en plus de remplir différentes fonctions discursives, les BDU ne se distribuent pas au hasard selon le genre de discours qui est en jeu. Les auteures ont dès lors pu montrer que les BDU remplissent une fonction cognitive dans la planification et l'interprétation du discours, et qu'un genre peut se caractériser selon le type de BDU qu'il contient. Nous corroborons ce point au sein de la section 4.1, où nous étudions la distribution des BDU au sein d'un plus large corpus, présenté dans la section 2.

Parmi les quatre types de BDU existants, les BDU régulatrices constituent un cas particulier : constituées uniquement d'un élément syntaxiquement non régi tel qu'un connecteur ou un marqueur du discours, elles se caractérisent par l'absence de contenu sémantique prédicatif. Pour Chafe, ces unités régulatrices remplissent un rôle bien particulier pour la structuration du discours :

Some regulate the development of the discourse [...]. Others have to do with the interaction between the participants [...]. Still others express the speaker's mental process, or judgment of the validity of the information being conveyed.

(Chafe, 1994 :64)

Le corpus étudié par Degand & Simon (2009b) ayant fait apparaître la majorité de ses unités régulatrices (68 %) en *périphérie gauche* de l'énoncé, nous pensons que cette position particulière joue un rôle stratégique dans la structuration des discours. La périphérie gauche constitue en effet le point d'ancrage du message, où le contenu du discours peut être relié de manière cohérente aux contenus précédents (Virtanen, 2004 : 80-81) et joue un rôle dans l'organisation informationnelle du discours. Selon Ho-Dac (2007), cette position (initiale) est stratégique pour le développement du discours. À travers cette étude, nous poursuivons dès lors l'hypothèse selon laquelle les éléments périphériques constituent des expressions cohésives au service de la cohérence discursive, et qu'ils participent d'une stratégie discursive. Nous pensons également qu'à l'instar des BDU, la forme et la fonction des éléments en périphérie gauche varient avec la situation de production du discours. Afin d'étayer ces hypothèses, nous explicitons en section 3 notre définition de la périphérie gauche et étudions, en section 4, sa distribution (ainsi que celle des BDU) au sein de notre corpus, au niveau de ses formes et de ses fonctions.

1. La périphérie gauche : études antérieures

De nombreux auteurs se sont intéressés à la notion de périphérie gauche et ce, sous des angles théoriques variés. La majorité des études se concentrent préférentiellement sur les constructions dites à détachement, que Marandin (1999) définit comme des constituants incidents, car « extérieurs au contenu de l'énoncé » (Dessaintes, 1960 : 13, cité par Marandin, 1999). Pour Apothéloz, Combettes & Neveu (2009), le principal point commun de ces constructions détachées « est que leur logique se situe à l'interface de la syntaxe et du discours : phrases dites *disloquées* ou *segmentées* (...), appositions, appendices divers et autres phénomènes de discontinuité syntaxique » (Apothéloz, Combettes & Neveu, 2009 : 1). Pour les auteurs, ces constructions interviennent souvent dans la problématique de la cohésion du discours et de l'organisation de l'information, mais aussi dans celle de la valeur interactionnelle des énoncés et des propriétés prosodiques des énonciations.

Devant la diversité des études existantes sur les éléments situés en périphérie gauche, nous dégageons trois approches principales (selon Degand, Simon, Tanguy & Van Damme, à paraître) :

- les approches positionnelles, qui étudient les rôles discursifs des éléments situés en position initiale d'énoncé (Fries, 1995 ; Marandin, 1999 ; Virtanen, 1992 et 2004), mais aussi des textes de manière plus générale (Ho-Dac, 2007), en problématisant la question de la linéarisation ;
- les approches (macro-)syntaxiques, qui traitent de l'organisation interne des énoncés à travers les incidences prosodiques, sémantiques et pragmatiques des constructions détachées à gauche (voir contributions dans Apothéloz *et al.*, 2009) ;
- les approches organisationnelles, qui considèrent les constructions périphériques – en particulier les adverbiaux – comme un indice de l'organisation du discours (Diessel, 2005) et, plus spécifiquement, comme des introducteurs de nouveaux cadres discursifs (Charolles, 1997 ; Charolles *et al.*, 2005 ; Combettes, 2005 ; Le Draoulec & Péry-Woodley, 2005), pouvant assumer des fonctions à l'échelle du discours (Charolles & Lamiroy, 2012).

Notre conception de la périphérie gauche se situe à la jonction de ces trois axes de recherche tout en s'en démarquant, puisque nous ne nous limitons pas à une seule position ou fonction (cf. Virtanen, 2004 et Charolles *et al.*, 2005 qui traitent isolément les adverbes de temps, de lieu et de manière en position initiale), ni à une liste fermée de catégories syntaxiques (voir contributions dans Apothéloz *et al.*, 2009, qui traitent uniquement de nombreuses formes de dislocation syntaxique). Les critères que nous utilisons pour la définition de la périphérie gauche sont en effet plus larges : nous les détaillons en section 3, après la présentation, en section suivante, du corpus sur lequel nous basons nos recherches.

2. Le corpus LOCAS-F

Les tableaux 1 et 2 (en annexes) offrent une description générale de notre corpus LOCAS-F (pour *Louvain Corpus of Annotated Speech – French* ; Degand, Martin & Simon, 2014).

Celui-ci est constitué de 42 échantillons sonores, répartis en 12 « genres » ou activités de communication orale différentes (nous suivons ici la conception du genre comme pratique sociale, telle que préconisée par Miller en 1984 ou Rastier en 1989). Chaque genre est représenté par 3 à 5 enregistrements différents, d'une durée de 3 à 5 minutes chacun. Le corpus totalise 48 locuteurs, représentant principalement des variétés du français de Belgique et de France.

Afin de rendre les données comparables entre elles et de pouvoir caractériser chaque enregistrement, nous avons catégorisé ceux-ci selon trois traits situationnels, qui permettent de « catégoriser le comportement communicatif des interlocuteurs par rapport aux déterminants situationnels et contextuels » (Koch & Oesterreicher, 2001 : 586) et qui sont susceptibles, selon nous, d'influencer la manière dont le discours est structuré² :

- le degré d'interactivité entre les participants : 0 désigne un discours non interactif ; 1 un discours semi-interactif (liberté d'interrompre restreinte) ; 2 un discours interactif (parole librement distribuée) ;

- le degré de préparation du discours : 0 caractérise un discours spontané ; 1 un discours semi-préparé (le sujet est connu par le participant) ; 2 un discours préparé ou lu ;

- le caractère médiatique du discours : 0 désigne des données non médiatiques ; 1 des données secondairement médiatiques (où l'activité implique plusieurs rôles communicatifs, les participants ne s'adressant pas directement aux auditeurs/spectateurs) ; 2 des données médiatiques, exclusivement produites pour être diffusées.

Le corpus LOCAS-F est annoté au niveau syntaxique (unités de rection et séquences fonctionnelles) et au niveau prosodique (unités prosodiques majeures et intermédiaires). De la mise en correspondance entre l'annotation rectionnelle et en unités prosodiques majeures résulte la segmentation en BDU et

² Cette sélection de trois critères situationnels est effectuée en regard de l'étude de Simon *et al.* (2010) selon laquelle l'ensemble des traits proposés par Koch & Oesterreicher (2001) peuvent se réduire à ces trois dimensions pertinentes lorsqu'il s'agit de refléter la variation stylistique d'échantillon de parole et de définir des phonogenres.

l'identification de leur périphérie gauche, toutes deux réalisées de manière automatique.

3. La périphérie gauche : notre définition

Puisque nous considérons les marques syntaxiques et les marques prosodiques comme équivalentes dans leur aptitude à délimiter des unités de discours, nous adoptons une méthode similaire, pour l'identification des périphéries gauches, à celle utilisée pour les BDU. Ainsi, nous identifions les éléments en périphérie gauche grâce à la combinaison de marques syntaxiques et prosodiques : un élément peut être considéré comme périphérique dès lors qu'il est isolé par une frontière syntaxique (élément non régi) ou par une frontière prosodique majeure, ou encore par une combinaison de ces deux marques. Lorsque l'initiale de la BDU est intégrée syntaxiquement *et* prosodiquement au reste de la séquence, nous l'identifions comme un simple élément initial. La combinaison de ces deux critères syntaxique et prosodique donne lieu à quatre types principaux (Degand, Simon, Tanguy & Van Damme, à paraître) :

a) *EI* – Élément initial : l'élément est prosodiquement et syntaxiquement intégré à la BDU. Exemple³ :

[(**nous sommes**)_{SV} (je vous l'ai dit tout à l'heure)_{insert} (treize mois avant l'élection)_{SO}]^{urv+} [(notre travail)_{SS} (n'est pas terminé)_{SV}]^{urv} /// (BDU-I)

b) *EPG-S* – Élément périphérique gauche isolé par une frontière syntaxique. Exemple :

<**mais**>_{md} [(on travaille)_{SV} (sur une représentation stylisée de la fréquence fondamentale)_{SO}]^{urv} /// (BDU-C)

c) *EPG-I* – Élément périphérique gauche isolé par une frontière prosodique. Exemple :

[(**l'homme** /// se rendit compte)_{SV} /// (qu'il était sur son passage et que la jeune femme ralentissait pour ne pas le heurter)_{SO}]^{urv} /// (BDU-S)

³ Les séquences encadrées par [...] correspondent à des unités de rection, celles encadrées par (...) à des séquences fonctionnelles (séquence sujet, séquence verbale, séquence objet, etc.). Les chevrons <...> marquent les éléments non rectionnels. Le symbole /// indique une frontière prosodique majeure.

d) *EPG* – Élément périphérique gauche isolé par une frontière syntaxique *et* prosodique (correspond à une BDU régulative).
Exemple :

<honnêtement>_{ag} /// (BDU-R)
[(je crois)_{SV} (que c'est pour faire bon garçon respectable)_{SO}]^{urv}
/// (BDU-C)

Avec notre méthode d'identification des éléments périphériques, deux autres types d'éléments ont émergé de nos données. Il s'agit tout d'abord des éléments périphériques complexes qui comportent un élément régulateur non réactionnel (marqueur de discours, connecteur...) ainsi qu'une partie de l'unité de réaction (et de la première séquence fonctionnelle) qui, elle, est isolée par la prosodie. Nous les avons étiquetés « EPG-S/I », dont voici un exemple ci-dessous :

<mais>_{md} [(si tu aimes ton frère /// et que tu n'aimes pas Dieu)_{SRg} /// (est-ce que)_{mi} (très vite)_{SRg} ton amour (ne sera-t-il pas)_{SV} (un peu /// rétréci)_{SO}]^{urv+} /// (BDU-S)

Nous traitons ce cas de manière particulière, car ces périphéries encodent deux contenus (conceptuels et/ou procéduraux) distincts. Notre méthode de codage fonctionnel ne distinguant pas actuellement ce cas de « double périphérie » des périphéries gauches simples, nous laissons ces EPG-S/I de côté pour la présente analyse fonctionnelle (section 4.2.2).

D'autre part, de nombreuses BDU sont apparues comme ne contenant qu'une seule séquence fonctionnelle (souvent verbale) et une seule unité prosodique, comme dans l'exemple suivant :

[(je peux te tutoyer)_{SV}]^{urv} (BDU-C)

De par leur structure, il n'est pas possible de délimiter une périphérie gauche à ces BDU. Étiquetées « MONO », nous les avons dès lors écartées de notre analyse formelle et fonctionnelle des périphéries gauches.

4. Analyses

4.1 Distribution des BDU

Afin de corroborer l'étude de Degand & Simon (2009b) selon laquelle les BDU ne se distribuent pas au hasard en fonction des quatre genres que les auteurs ont étudiés, nous étudions ici la distribution des 2 347 BDU de notre corpus à travers les 12 activités de communication qui le constituent. La figure 2 ci-dessous confirme que les types de BDU varient selon le genre en jeu, et que leur distribution n'est pas aléatoire. Cette différence de distribution est en effet statistiquement significative ($\chi^2 = 602,84$; $dl = 44$; $p < 0,001$).

Le calcul des résidus de la statistique de Chi-carré nous permet de constater que les discours académiques, les discours politiques et les homélies se démarquent des autres genres par leur proportion significativement élevée de BDU regroupées par la syntaxe. C'est-à-dire que la parole y est comme « sursegmentée » intonativement. Pour Degand & Simon (2011), cela dénote un style de parole didactique ou emphatique, qui est typique des genres où le degré de préparation est élevé. On remarque en effet que les genres qui comprennent un nombre élevé de BDU regroupées par la syntaxe (partie haute du graphique en figure 2) sont caractérisés par un degré de préparation élevé, mais aussi par une non-interactivité et par une certaine médiatisation (la diffusion étant limitée pour le discours académique).

Inversement, les genres caractérisés par une plus grande spontanéité (partie basse du graphique en figure 2) présentent une proportion plus élevée de BDU groupées par l'intonation. Ainsi, les narrations conversationnelles et l'interview libre se distinguent des autres genres en délivrant davantage la parole en grands « paquets informatifs ». Une telle stratégie discursive permet selon nous de marquer des liens de cohérence entre plusieurs unités syntaxiques, sans pour autant les marquer de manière explicite.

Les BDU congruentes et les BDU régulatrices se distribuent davantage de manière homogène au sein du corpus, excepté pour les interviews libres et les discours académiques, qui présentent une proportion significativement élevée de BDU congruentes et de BDU régulatrices (respectivement). Les BDU

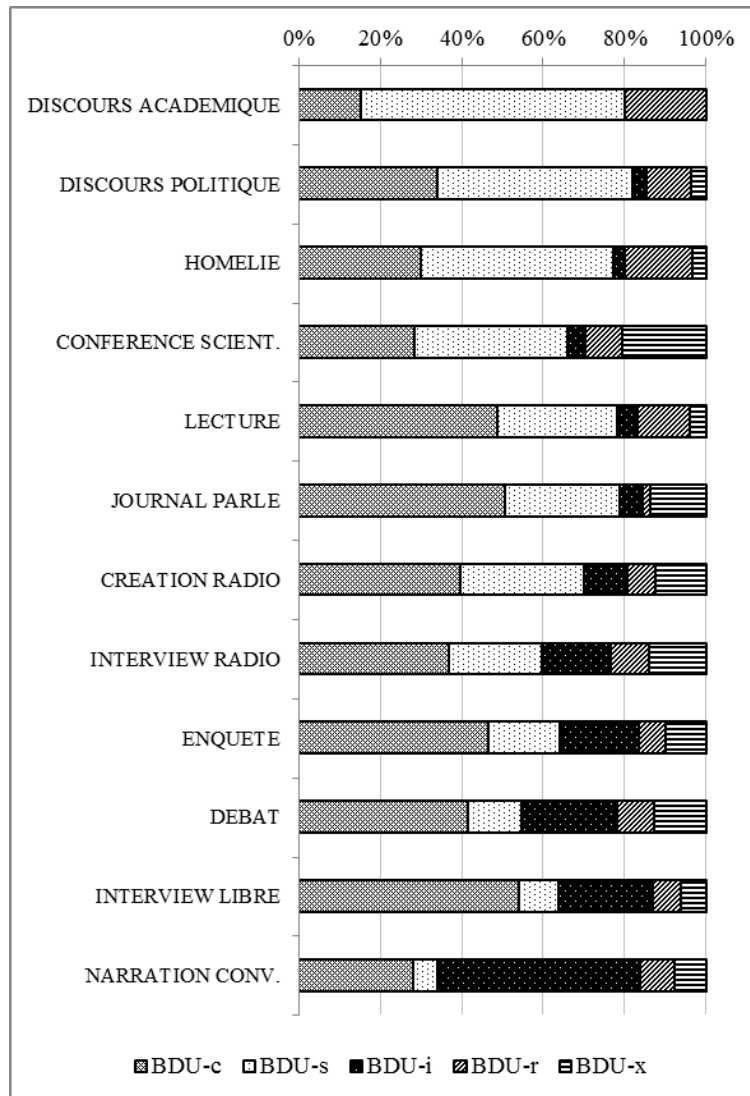


Figure 2. Distribution des BDU à travers les différents genres du corpus LOCAS-F

congruentes indiquent, selon nous, une présentation de l'information de manière neutre, tandis que les BDU régulatrices remplissent *a priori* des fonctions textuelles, des relations de discours et de gestion des topiques. Ces BDU régulatrices se trouvant majoritairement en périphérie gauche d'énoncé (cf. supra), nous présentons dans la section suivante l'étude de leurs distributions formelle et fonctionnelle, ainsi que celles des autres types d'éléments périphériques.

4.2 Distributions formelle et fonctionnelle des EPG

4.2.1 Distribution formelle

La figure 3 ci-dessous nous permet de constater qu'à l'instar des BDU, les éléments périphériques à gauche et éléments initiaux ne se distribuent pas au hasard en fonction des différents genres du corpus ($\chi^2 = 622,84$; $dl = 55$; $p < 0,001$). Plus particulièrement, la distribution des EPG-S et des EPG-I s'inverse au fur et à mesure du graphique : la proportion d'EPG isolés par la syntaxe est significativement élevée dans les narrations conversationnelles et les créations radio, et a tendance à fortement baisser lorsqu'il s'agit de genres où le degré de préparation est plus élevé. Ce paramètre semble en effet influencer la proportion d'EPG-S et d'EPG-I au sein du genre représenté. À l'inverse, la proportion d'EPG isolés par la prosodie est significativement basse pour ces mêmes genres situés dans la partie haute du graphique, et grandit lorsqu'on se dirige vers la partie basse du graphique, pour devenir significativement élevée au sein des discours académiques, des discours politiques et des homélies. Nous l'avons vu dans la section précédente, ces trois genres sont également caractérisés par une proportion significativement élevée de BDU groupées par la syntaxe. Une telle proportion d'EPG-I semble dès lors être le reflet de la « sursegmentation » prosodique dont ces genres font preuve.

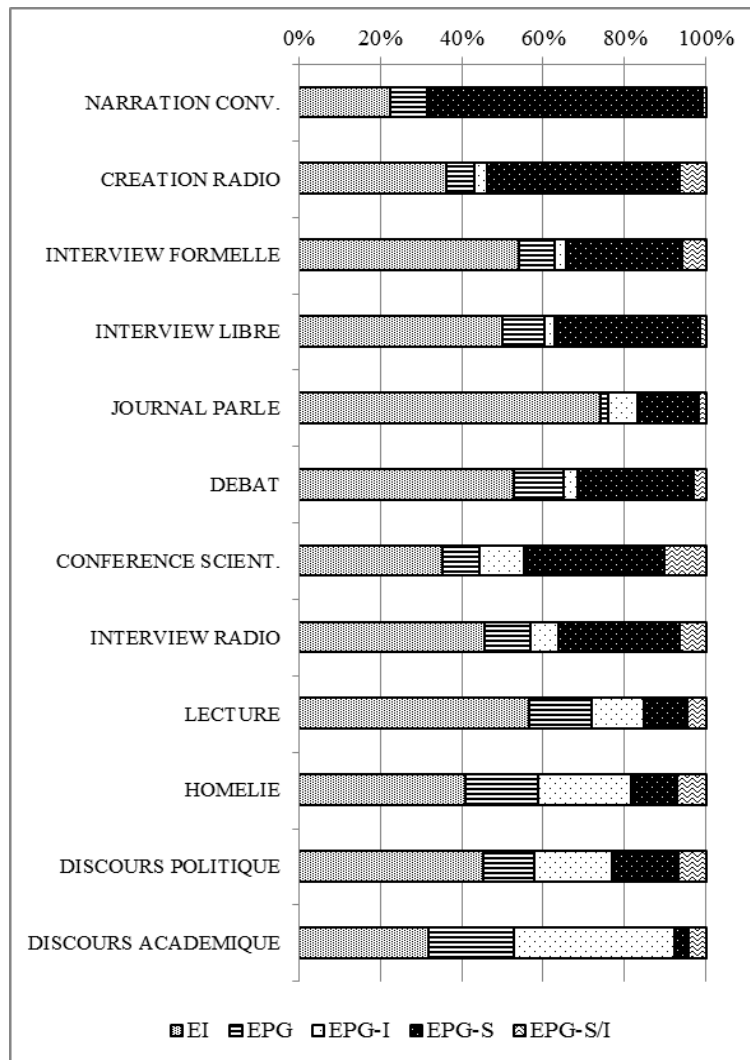


Figure 3. Distribution des éléments à gauche (forme) à travers les différents genres du corpus LOCAS-F

4.2.2 Distribution fonctionnelle

Catégories fonctionnelles

Les catégories choisies pour l'étiquetage fonctionnel des périphéries gauches sont inspirées de Chafe (1994 : 64). Celles-ci se répartissent en deux niveaux d'annotation. Le premier distingue la fonction textuelle, orientée vers le texte et l'organisation discursive, la fonction interactive, orientée vers les participants, et la fonction subjective, orientée vers le locuteur.

Ainsi, lorsqu'il revêt une fonction textuelle, l'élément situé en périphérie gauche introduit une relation sémantique entre deux contenus propositionnels du discours, en marquant soit une relation de cohérence, soit une marque de gestion topicale, ou encore une relation métadiscursive. Lorsqu'il encode une fonction interactive, l'élément périphérique marque soit une interpellation lorsque le locuteur attire l'attention de son interlocuteur dans son discours, soit un changement de tour de parole lorsqu'il signale la prise de parole du locuteur, ou encore un backchannel lorsqu'il indique une confirmation de la part de l'interlocuteur, sans volonté de prise de parole de sa part. Enfin, lorsqu'il marque une fonction subjective, l'élément périphérique marque le jugement de validité ou le processus mental du locuteur, ou fait simplement référence à lui-même, à travers une fonction de deixis. Le tableau 3 présenté en annexe donne un exemple pour chaque catégorie.

Accord inter-annotateur

L'annotation fonctionnelle a été réalisée par deux annotateurs, 1/3 du corpus ayant été doublement annoté pour pouvoir calculer un score inter-annotateur (Kappa score). L'annotation fut d'abord réalisée sur toutes les « unités périphériques » du corpus, éléments initiaux (EI) et MONO compris. Il est rapidement apparu que les catégories fonctionnelles sélectionnées étaient opérationnelles uniquement pour les éléments réellement périphériques : nous avons noté un haut niveau d'inconfort lorsqu'il s'agissait d'attribuer une catégorie à un élément EI ou MONO, qui ne forment pas vraiment de périphérie gauche sur leur plan formel. Le Kappa score calculé sur la totalité de nos annotations, EI et MONO compris, est par conséquent très bas (0.34). Lorsque nous retirons ces « fausses périphéries » du

calcul, nous obtenons un Kappa score beaucoup plus élevé, de 0.87. Ce score indique, d'une part, que nos catégories fonctionnelles sont bien opérationnalisées et distinctes les unes des autres et, d'autre part, que les EI et les MONO sont fonctionnellement distincts des autres périphéries identifiées. Nous les écartons dès lors de nos analyses pour le moment.

Analyse

La majorité des éléments situés en périphérie gauche (80 %) participent à la progression textuelle via des relations de cohérence surtout (52,6 %), puis des relations topicales (22,8 %) et métadiscursives (4,2 %). Seuls 11 % des éléments périphériques remplissent une fonction interactive, avec 6,6 % de marques de turn-taking, 4,1 % d'interpellations et 0,4 % de back-channel. Une grande partie des éléments de type « MONO » étiquetés (et supprimés de la présente analyse) remplissent cette fonction, avec de nombreuses BDU de type [(ouais)_{SAdv}]^{ura}/// ou [(oui oui)_{SAdv}]^{ura}///, qui traduisent une confirmation de la part de l'interlocuteur. Enfin, 9 % des périphéries gauches remplissent une fonction subjective, dont 6,8 % expriment le point de vue du locuteur, et 2,5 % une fonction de deixis.

Tableau 1. Répartition des catégories fonctionnelles selon la forme des périphéries gauches

Fonction	Sous-Fonction	EPG	EPG-S	EPG-I
Textuelle	<i>Cohérence</i>	<u>82</u>	366	<u>49</u>
	<i>Topic</i>	43	<u>57</u>	115
	<i>Métadiscours</i>	11	27	2
Interactive	<i>Turn-taking</i>	20	42	<u>0</u>
	<i>Interpellation</i>	25	<u>10</u>	4
	<i>Backchannel</i>	4	<u>0</u>	0
Subjective	<i>Point de vue</i>	16	36	12
	<i>Déixis</i>	<u>0</u>	12	12

Le tableau 1 ci-dessus montre la répartition des éléments périphériques selon la fonction qu'ils occupent. Statistiquement, les

fonctions ($\chi^2 = 54,66$; $dl = 4$; $p < 0,001$) et les sous-fonctions ($\chi^2 = 301,16$; $dl = 14$; $p < 0,001$) des éléments périphériques se distribuent différemment selon le type de frontière (prosodique, syntaxique ou une combinaison des deux) qui les détache du reste de l'énoncé. Les sous-fonctions expliquent néanmoins mieux cette variation de distribution (V de Cramer⁴ = 0,40) que les fonctions seules (V de Cramer = 0,17). Une analyse des résidus de Pearson associés⁵ permet une analyse plus fine de cette variation. Nous observons que les EPG (détachement syntaxique et prosodique) remplissent préférentiellement des fonctions interactives (résidus significatifs pour les trois sous-fonctions de turn-taking, d'interpellation et de backchannel) au détriment de la fonction textuelle de cohérence, tandis que les EPG-S (détachement syntaxique uniquement) se caractérisent par une surreprésentation de cette fonction de cohérence et une sous-représentation de la gestion topicale. À l'inverse, les EPG-I (détachement prosodique) marquent plus fréquemment une fonction topicale ou de deixis, et significativement peu de fonctions textuelle de cohérence.

Afin de déterminer si les fonctions des éléments situés en périphérie gauche traduisent différentes stratégies discursives, il s'agit d'étudier leur distribution à travers les différents genres du corpus LOCAS-F. La mise en relation de ces deux variables ne donnant pas de résultat satisfaisant (trop grand nombre de faibles effectifs théoriques, calcul du Chi-Carré impossible), nous regroupons ici les genres du corpus selon leurs traits situationnels. Plus particulièrement, nous étudions la distribution des catégories fonctionnelles selon le degré d'interactivité du genre en question (voir tableau 5 ci-dessous). La statistique de Chi-Carré confirme en effet que ces deux variables ne sont pas indépendantes ($\chi^2 = 156,68$; $dl = 14$; $p < 0,001$).

⁴ Mesure d'association permettant de comparer des valeurs de Chi-Carré entre elles et de déterminer la force de la relation entre les deux variables testées.

⁵ Les résidus mesurent la contribution de chaque observation à la significativité du test de statistique. Dans nos tableaux, les observations en gras sont significativement plus fréquentes qu'attendues, alors que les observations soulignées sont significativement moins fréquentes qu'attendues.

Tableau 2. Répartition des catégories fonctionnelles selon le degré d'interactivité du discours

Fonction	Sous-Fonction	Interacti-vité 0	Interacti-vité 1	Interacti-vité 2
Textuelle	<i>Cohérence</i>	304	89	<u>104</u>
	<i>Topic</i>	150	30	<u>35</u>
	<i>Métadiscours</i>	17	8	15
Interactive	<i>Turn-taking</i>	<u>0</u>	13	49
	<i>Interpellation</i>	19	5	15
	<i>Backchannel</i>	<u>0</u>	0	4
Subjective	<i>Point de vue</i>	<u>22</u>	12	30
	<i>Déixis</i>	10	6	8

Les genres de discours non-interactifs (tels que le discours politique, la conférence scientifique ou le journal parlé), que nous pouvons placer du côté de la « distance communicative » (Koch & Oesterreicher, 2001) se caractérisent par une sur-représentation des éléments périphériques gauches à fonction de cohérence et de gestion des topiques ainsi que par une sous-représentation des fonctions de gestion de tour de parole ou qui expriment un point de vue. De ce fait, ces genres se définissent davantage par des stratégies discursives de planification et d'organisation du discours. Dans les genres caractérisés par un degré d'interactivité plus élevé, tels que le débat ou l'interview libre, la tendance s'inverse : les périphéries gauches propres à ces genres sont moins souvent textuelles, et remplissent préférentiellement des fonctions interactives de turn-taking ainsi que des fonctions subjectives en exprimant le point de vue du locuteur. Ces deux dernières fonctions caractérisent en effet davantage les genres proches de l'« immédiat communicatif », où il s'agit de mettre en œuvre des stratégies de gestion et de négociation du flux de parole conversationnelle, et de marquer la subjectivité du locuteur.

Conclusion

Dans cette contribution, nous proposons une extension de la notion de périphérie gauche en définissant cette dernière en termes syntaxiques et prosodiques, à l'instar des BDU, qui constituent l'unité de segmentation de notre corpus.

Nous avons montré que la distribution des BDU ainsi que celle des périphéries gauches (forme et fonction) varient en fonction du genre qui est en jeu et de ses traits situationnels. Le « profil » en BDU et en périphérie gauche d'un genre est motivé par des stratégies discursives différentes.

En effet, les genres représentatifs de la distance communicative, avec un degré de préparation élevé et une absence d'interaction entre les locuteurs, présentent davantage de BDU groupées par la syntaxe et d'éléments périphériques isolés à gauche par la prosodie, marquant ainsi des stratégies discursives didactiques ou emphatiques. Les périphéries gauches représentées au sein de ces genres remplissent préférentiellement des fonctions textuelles, de cohérence et de gestion des topiques, cette dernière fonction étant d'ailleurs celle dans laquelle se « spécialise » l'EPG-I. À l'inverse, les genres représentatifs de l'immédiat communicatif privilégient les BDU groupées par l'intonation ainsi que des éléments périphériques qui fonctionnent davantage comme marqueurs interactifs et subjectifs.

Ainsi, nous avons montré que les BDU et les éléments situés en périphérie gauche sont au service de stratégies discursives définies. Il reste à analyser de manière approfondie les éléments initiaux et autres cas particuliers, laissés de côté pour le moment. Nous souhaitons également nous attarder plus précisément sur les BDU régulatrices (périphéries isolées par une frontière syntaxique et une frontière prosodique), qui nous semblent primordiales dans la négociation de la structure discursive.

Références bibliographiques

- Apothéloz D., Combettes B. & Neveu F. (éd.) (2009). *Les linguistiques du détachement*. Berne : Peter Lang.
- Bilger M. & Campione E. (2002). « Propositions pour un étiquetage en 'séquences fonctionnelles' », *Recherches sur le français parlé* 17 : 117-136.

- Blanche-Benveniste C., Bilger M., Rouget Ch., Van Den Eynde K. & Mertens P. (1990). *Le français parlé : études grammaticales*. Paris : Editions du CNRS.
- Chafe W.L. (1994). *Discourse, consciousness, and time*. Chicago : University of Chicago Press.
- Charolles M. (1997). « L'encadrement du discours : univers, champs, domaines et espaces », *Cahier de Recherche Linguistique*, Landisco, Université Nancy 2, 6 : 1-73.
- Charolles M., Le Draoulec A., Pery-Woodley M-P. & Sarda L. (2005). « Temporal and spatial dimensions of discourse organization », *French Language Studies* 15 : 115-130.
- Charolles M. & Lamiroy B. (2012). « Les adverbiaux : de la phrase au discours », *31^e Colloque International sur le Lexique et la Grammaire*, 93-99.
- Combettes B. (2005). « Les constructions détachées comme cadres de discours », *Langue française* 148 : 31-44.
- Degand L., Martin L.J. & Simon A-C. (2014). « Unités discursives de base et leur périphérie gauche dans LOCAS-F, un corpus oral multigenres annoté », SHS Web of Conferences 8, 2613-2625 [Accessible en ligne : http://www.shs-conferences.org/articles/shsconf/abs/2014/05/shsconf_cmlf14_01211/shsconf_cmlf14_01211.html#].
- Degand L. & Simon A-C. (2008). « Minimal Discourse Units in Spoken French : Un-covering Genre-bound Segmentation Strategies », in W. Ramm et C. Fabricius-Hansen (éd.) *Linearisation and Segmentation in Discourse*. Lysebu, Oslo : Department of Literature, Area Studies and European Languages, University of Oslo, 31-42.
- Degand L. & Simon A-C. (2009a). « On identifying basic discourse units in speech : theoretical and empirical issues », *Discours* 4 [Accessible en ligne : <http://discours.revues.org/index.html>].
- Degand L. & Simon A-C. (2009b). « Mapping prosody and syntax as discourse strategies : How Basic Discourse Units vary across genres », in A. Wichmann, D. Barth-Weingarten et N. Dehé (éd.) *Where prosody meets pragmatics : research*

- at the interface*. Studies in Pragmatics. Bingley : Emerald, 79-105.
- Degand L. & Simon A-C. (2011). « L'analyse en unités discursives de base : pourquoi et comment ? », *Langue française* 170 : 45-59.
- Degand L., Simon A-C, Tanguy N. & Van Damme, Th. (à paraître). « Initiating a discourse unit in spoken French : Prosodic and syntactic features of the left periphery », in S. Pons Borderia (éd.) *Theoretical Models of Discourse Segmentation*. Amsterdam : John Benjamins.
- Dessaintes M. (1960). *La construction par insertion incidente*. Paris : D'Artrey.
- Diessel H. (2005). « Competing motivations for the ordering of main and adverbial clauses », *Linguistics* 43/3 : 449-470.
- Fries P-H. (1995). « Patterns of information in initial position in English », in P-H Fries & M. Gregory (éd.) *Discourse in society : systemic functional perspectives*, 47-66.
- Halford B. K. (1996). *Talk Units : The Structure of Spoken Canadian English*. Tübingen : Narr.
- Ho-Dac L-M. (2007). *La position initiale dans l'organisation dans l'organisation du discours : une exploration en corpus*. Thèse de doctorat, Toulouse : Université Toulouse le Mirail.
- Koch P. & Oesterreicher W. (2001). « Langage parlé et langage écrit », in G. Holtus, M. Metzeltin & C. Schmitt (éd.) *Lexikon der romanistischen Linguistik* 1/2 : 584-627.
- Lacheret-Dujour A. & Victorri B. (2002). « La période intonative comme unité d'analyse pour l'étude du français parlé : modélisation prosodique et enjeux linguistiques », *Verbum* 24 : 55-72.
- Le Draoulec A. & Pery-Woodley M-P. (2005). « Encadrement temporel et relations de discours », *Langue française* 148 : 45-60.
- Marandin J-M. (1999). *Grammaire de l'incidence*. [Accessible en ligne : <http://www.llf.cnrs.fr/Gens/Marandin/archive-fr.php>].

- Mertens P. & Simon A-C. (2013). « Towards automatic detection of prosodic boundaries in spoken French », in P. Mertens & A-C. Simon (éd) *Proceedings of the Prosody-Discourse Interface Conference 2013 (IDP-2013)* : 81-87.
- Miller C-R. (1984). « Genre as social action ». *Quarterly journal of speech*, 70/2 : 151-167.
- Rastier F. (1989). *Sens et Textualité*. Paris : Hachette.
- Roulet E. (1991). « A propos des niveaux de l'analyse linguistique », in R. Liver, I. Werlen & P. Wunderli (éd.) *Sprachtheorie und Theorie der Sprachwissenschaft*. Tübingen : Gunter Narr Verlag, 221-230.
- Selting M. (2000). « The construction of units in conversational talk », *Language in Society* 29 : 477-517.
- Simon A-C., Auchlin A., Avanzi M. & Goldman J-Ph. (2010). « Les phonostyles. Une description prosodique des styles de parole en français », in M. Abecassis & G. Ledegen (éd.) *Les voix des Français. En parlant, en écrivant*. Berne : Peter Lang, 2 : 71-88.
- Virtanen T. (1992). *Discourse Functions of Adverbial Placement in English*. Bo : Bo Akademi University Press.
- Virtanen T. (2004). « Point of departure : Cognitive aspects of sentence-initial adverbials », in T. Virtanen (éd.) *Approaches to cognition through text and discourse*. Berlin/ New York : Mouton de Gruyter, 79-97.

Annexes

Tableau 1. Description générale du corpus (a)

Genre	Durée (min:sec)	Degré d'interactivité	Degré de préparation	Caractère médiatique
<i>Conférence scientifique</i>	16:44	0	2	0
<i>Débat</i>	19:19	2	1	1
<i>Discours académique</i>	15:18	0	2	1
<i>Discours politique</i>	20:26	0	2	2
<i>Homélie</i>	13:23	0	2	2
<i>Enquête</i>	16:04	1	1	2
<i>Interview libre¹</i>	15:33	2	0	1
<i>Interview radio</i>	20:29	1	1	1
<i>Journal parlé</i>	14:45	0	2	2
<i>Lecture</i>	15:18	0	2	1
<i>Narration conversationnelle</i>	10:22	0	0	0
<i>Création radio²</i>	13:24	0	1	2
Total	3:11:05			

¹ Interview de type informel, réalisée dans le cadre de l'émission belge "Hep Taxi!".

² Temps d'un 45 tours devant un micro ouvert, cfr. www.ateradio.com, classé sous "chroniques - art".

Tableau 2. Description générale du corpus (b)

Genre	Nombre d'unités syntaxiques	Nombre d'unités prosodiques	Nombre de BDU	Nombre de mots	Moyenne mots/bdu
<i>Conférence scientifique</i>	140	320	112	3123	27,88
<i>Débat</i>	535	528	362	4790	13,23
<i>Discours académique</i>	100	401	125	2508	20,06
<i>Discours politique</i>	194	475	194	3300	17,01
<i>Homélie</i>	136	343	150	1902	12,68
<i>Enquête</i>	292	334	221	2575	11,65
<i>Interview libre</i>	500	476	370	3559	9,62
<i>Interview radio</i>	358	474	258	4288	16,62
<i>Journal parlé</i>	162	206	124	2938	23,69
<i>Lecture</i>	176	414	184	3157	17,16
<i>Narration conversationnelle</i>	264	133	117	2320	19,83
<i>Création radio</i>	168	277	131	2452	18,72
Total	3035	4381	2348	36912	17,35

Tableau 3. Catégories fonctionnelles pour les périphéries gauches

Fonction	Sous-fonction	Exemple
Textuelle	Cohérence	<donc> _{md} [(nous y sommes // habitués) _{sv}] _{lurv} ///
	Gestion topicale	<les sentiments> _{ag} [(par moment) _{SRg} (ça me gonfle un peu) _{sv}] _{lurv} ///
	Métadiscursive	[(je vais finir) _{sv}] (après avoir évoqué ces points) _{SRd} (en vous parlant de la paix) _{SRd}] _{lurv} ///
	Interpellation	eah <madame Fressoz> _{ag} [(il y a) _{sv} (deux camps // en lutte) _{so}] (dans le monde entier) _{SRd}] _{lurv} ///
Interactive	Turn-taking	<mais> _{md} [(pensez-vous) _{sv} (que les accents régionaux soient une entrave à la compréhension) _{so}] _{lurv} ///
	Backchannel	L1 [(la faute) _{SN} (aux trente-cinq heures) _{SRd}] _{lura} [(vous rétorqueront-// ils) _{sv}] _{lurv} /// L2 <voilà> _{md} ///
	Point de vue	<honnêtement> _{ag} /// [(je crois) _{sv} (que c'est pour faire bon garçon respectable) _{so}] _{lurv} ///
Subjective	Déixis	<moi> _{ag} [(j'aime bien aussi) _{sv}] (lui donner l'impression qu'il est libre) _{lurv} ///